ARL-CR-0841 ● Nov 2019

DEVCOM ARMY RESEARCH LABORATORY

# M1901097 – Information Sciences Campaign, KCI-IS-1: Cyber Fire and Maneuver in Tactical Battle: Dynamic Watermarking and Timing Analysis to Protect Vehicular Cyber-Physical Systems

by PR Kumar, Swaminathan Gopalswamy, Riccardo Bettati, Lantian Shangguan, Kenny Chour, Bharadwaj Satchidanandan, Jaewon Kim, Woo Hyun Ko, and Gopal Kamath

under contract W911NF1920033

## NOTICES

### Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-CR-0841 ● Nov 2019


# M1901097 – Information Sciences Campaign, KCI-IS-1: Cyber Fire and Maneuver in Tactical Battle: Dynamic Watermarking and Timing Analysis to Protect Vehicular Cyber-Physical Systems

**PR Kumar, Swaminathan Gopalswamy, Riccardo Bettati, Lantian Shangguan, Kenny Chour, Bharadwaj Satchidanandan, Jaewon Kim, Woo Hyun Ko, and Gopal Kamath**
***Texas A&M Engineering Experiment Station***
***College Station, TX 77840-4030***

**under contract W911NF1920033**

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED (From - To) |
|---|---|---|
| November 2019 | Contractor Report | 1 December 2018–31 October 2019 |

**4. TITLE AND SUBTITLE**
M1901097 – Information Sciences Campaign, KCI-IS-1: Cyber Fire and Maneuver in Tactical Battle: Dynamic Watermarking and Timing Analysis to Protect Vehicular Cyber-Physical Systems

**5a. CONTRACT NUMBER**
W911NF1920033

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
PR Kumar, Swaminathan Gopalswamy, Riccardo Bettati, Lantian Shangguan, Kenny Chour, Bharadwaj Satchidanandan, Jaewon Kim, Woo Hyun Ko, and Gopal Kamath

**5d. PROJECT NUMBER**
0011264600

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Texas A&M Engineering Experiment Station
7607 Eastmark Drive
College Station TX 77840-4030

**8. PERFORMING ORGANIZATION REPORT NUMBER**
ARL-CR-0841

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
US Army Research Office
800 Park Office Drive, Suite 4229
Research Triangle Park NC 27709

**10. SPONSOR/MONITOR'S ACRONYM(S)**
ARO

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**
ORCID ID(s): PR Kumar, 0000-0003-0389-5367; Swaminathan Gopalswamy, 0000-0003-3150-4589; Riccardo Bettati, 0000-0002-5877-6667; Lantian Shangguan, 0000-0003-1527-3354; Kenny Chour, 0000-0002-3799-1667; Bharadwaj Satchidanandan, 0000-0003-1913-7368; Jaewon Kim, 0000-0002-3021-729X; Woo Hyun Ko, 0000-0001-6378-6301; Gopal Kamath, 0000-0002-3190-9087.

**14. ABSTRACT**
Autonomous vehicles (AVs) can potentially bring about great benefits, such as higher safety and mobility, and lower congestion and fuel consumption; however, studies have shown that AVs are vulnerable to malicious cyber attacks. Since the breakdown of an AV can be catastrophic (e.g., passengers injured/killed), it becomes imperative to deploy defense techniques to enhance the cybersecurity of AVs, which is the motivation of this study. Dynamic watermarking (DW), an active defense technique for cyber-physical systems (CPSs), is adopted herein. Implemented in a real AV for an autonomous driving operation, the DW technique is shown to be robust and sensitive in the early detection of different types of malicious cyber attacks on the sensing system.

**15. SUBJECT TERMS**
dynamic watermarking, autonomous vehicles, cybersecurity, sensing, control, active defense, cyber-physical system

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Stephen Raio |
| Unclassified | Unclassified | Unclassified | UU | 23 | 19b. TELEPHONE NUMBER (Include area code) 410-278-9276 |

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

## Contents

## List of Figures

# 1. Introduction

## 1.1 Background

Autonomous vehicles (AVs) have the potential to reduce traffic deaths by over 90% (saving 30,000 lives in the United States each year),[1] increase fuel economy (up to 39%), and increase congested traffic speed (up to 13%), while providing mobility for the young, the elderly, and the disabled.[2] These benefits all assume a significant penetration of AVs. Such adoption of AVs is highly dependent on their safety and security, and studies[3–6] have shown that AVs are vulnerable to a variety of malicious cyber attacks such as masquerade, replay, message modification, and denial of service (DoS). The breakdown of an AV, independent of what causes it, can be catastrophic, as the human passengers could be seriously injured or even killed. It is therefore imperative to deploy defense techniques to mitigate the vulnerability and enhance the cybersecurity of AVs. Doing this will likely influence the adoption rate of AVs, thereby enabling society to reap the aforementioned benefits. This is the motivation of this project.

Dynamic watermarking (DW) is an active defense technique for cyber-physical systems (CPSs). Satchidanandan et al.[7] describe the technique, which is adopted in this project for detecting cyber attacks on the sensing system of an AV. This technique is intended to be a general purpose defense of CPSs. It has been tested in a simulation of attacks on the automatic gain control in an electronic power system.[8] DW has also been implemented and shown to be experimentally effective in a laboratory process control system of two cascaded tanks.[9] It was also investigated in a laboratory setting using radio-controlled cars over a WiFi network.[10] This project investigates whether this technique will work in a real-world environment that includes road noise and all the nonlinearities and complexities of real cars. We study the performance of the technique under a variety of attack scenarios and quantify its performance with respect to the level of watermark signal and detection time.

The DW technique consists of controllers adding small, private, random signals called “dynamic watermarks” to their control commands. Implemented by the actuators, these signals pass through the physical plant and are picked up as part of the ambient “noise” by the sensors. When the sensors report the measurements to the controllers, simple statistical tests can be conducted to verify the authenticity of the reported measurements, detecting cyber attacks (if any) on the sensing system. In this project, the performance of this approach is studied on a research AV. The effectiveness, time to detection, robustness, and sensitivity are validated in proving-ground tests on an instrumented autonomous Lincoln MKZ.

## 1.2 Related Works

Over the last few decades, a variety of studies have approached the cybersecurity issue of CPSs from different perspectives. Motivated by the scenario in which the adversarial presence causes a loss of data packets in the communication between sensors and controllers, a large body of research has been devoted to developing estimation and control algorithms for systems with intermittent observations.[11–15] Other studies have focused on developing techniques to counter attacks, such as the well-known replay attack (i.e., the adversary records sensor measurements for a period of time and replays them to maintain the illusion of a normal operating condition). Correspondingly, a methodology to secure systems from replay attacks was presented in multiple studies.[16–19] Specifically, Mo and Sinopoli[16] and Mo et al.[18] introduced a technique, termed physical watermarking, wherein the controller commands the actuator to inject into the system a random component not known in advance to secure the system against such an attack. An $\chi_2$ detector was proposed and extended by Weerakkody et al.[19] to detect an adversary that was capable of reporting false measurements to the estimator. Satchidanandan and Kumar[20] showed that such a single test does not suffice and can be successfully attacked. In an early work, they showed that two tests are indeed necessary and sufficient.[7] The DW approach is explored in greater detail in two other papers by the same authors.[21,22] Other examples of different approaches to the cybersecurity issue are detecting false data injection by zero-dynamics attacks,[23] exploiting correlations between multiple sensor measurements to weed out malicious sensors,[24] and studying the evolution of a physical process under DoS attacks (which prevent the controller and actuator from receiving required data) and deception attacks (which cause an actuator to issue incorrect actuation signals).[25]

Specifically on AV cybersecurity, a number of studies have investigated common cyber attacks and possible countermeasures. However, rather than developing specific techniques, most studies only provided high-level guidelines and principles for defending an AV's cybersecurity. For example, one study[3] identified and classified cyber attacks on AVs and discussed mitigation strategies including using a multiagent system architecture, micro-kernel, redundancy, diversity, authentication, and so on. Another study[26] proposed an approach for aligning safety and security life cycles at an early development phase using the failure, attack, and countermeasure (FACT) graph to connect safety failures, security attacks, and the associated countermeasures. Based on system-of-systems principles, one study[27] evaluated a multihomed, multi-level intrusion detection system for AVs. Specifically focusing on an intelligent vehicle's cybersecurity in the United States, Onishi et al.[5] examined the vulnerabilities related to carry-in devices and Global Navigation Satellite System (GNSS) and vehicle-to-vehicle (V2V) communication;

and discussed countermeasures such as complete penetration testing and integration of an attestation process.

To the best of our knowledge, the results of this report demonstrate the first real implementation of an active defense against arbitrary cybersecurity attacks on AVs.

## 2. Materials and Methods

### 2.1 Dynamic Watermarking Mechanism

The mechanism of the DW technique can be simply stated as follows. An actuator deliberately superimposes certain probing signals that are not disclosed to other nodes (i.e., sensors, controllers, other actuators) on top of the control-law specified signals. These probing signals are expected to appear in appropriately transformed ways at various points in the system. The controller can then conduct "tests" on the information received from the sensors to potentially infer whether there are malicious activities anywhere in the feedback path from the sensor measurements to the controller.

To understand the problem in precise terms in a simple context, we focus on a single-input, single-output linear stochastic dynamical system with Gaussian noise. The system is described by

$$x[t + 1] = ax[t] + bu[t] + \omega[t + 1], \quad (1)$$

where $a$, $b$, $x[t]$, $u[t]$, and $\omega[t] \in \mathsf{R}$, with $\{\omega[t]\}$ being the zero-mean, independent and identically distributed (i.i.d.) Gaussian noise of variance $\sigma_\omega^2$. The actuator implements a control law $\{g_t\}$, specifically, $u[t] = g_t(x^t)$, where $x^t := (x[0];x[1];...;x[t])$. For this, the actuator relies on a sensor that measures $x[t]$. However, the sensor could be malicious and report measurements $z[t]$ that differ from the true readings $x[t]$. We consider an honest actuator that implements the control policy $\{g\}$, but adds a private excitation $\{e\}$ as a defense. Specifically, the actuator applies to the system the input $u[t] = g_t(z^t) + e[t]$. Note that the private excitation $e[\cdot]$ added is i.i.d. and Gaussian with a mean 0 and variance $\sigma_e^2$. Therefore, the system evolves in closed loop as

$$x[t + 1] = ax[t] + bgt(zt) + be[t] + \omega[t + 1]. \quad (2)$$

As the core of the DW technique, certain "tests" are performed (by the controller) to check if the sensor is malicious or not. Toward developing these tests, note that the actual sequence of states $\{x[t]\}$ of the system satisfies

$$x[t + 1] - ax[t] - bg_t(z^t) = be[t] + \omega[t + 1]. \quad (3)$$

Therefore, we have

$$\{x[t+1] - ax[t] - bg_t(z^t)\}_t \sim i.i.d.N(0, b^2\sigma_e^2 + \sigma_\omega^2) \tag{4}$$

and

$$\{x[t+1] - ax[t] - bg_t(z^t) - be[t]\}_t \sim i.i.d.N(0, \sigma_\omega^2). \tag{5}$$

If the sensor were truthfully reporting $z[t] \equiv x[t]$, the reported sequence $\{z[t]\}$ would satisfy the conditions of Eqs. 4 and 5. The tests are formulated in an asymptotic form (over an infinite time interval) as follows and can be reduced to statistical tests over a finite time interval in standard ways:

- Test 1:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} (z[t+1] - az[t] - bg_k(z^k) - be[k])^2 = \sigma_\omega^2 \tag{6}$$

- Test 2:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} (z[t+1] - az[t] - bg_k(z^k))^2 = b^2\sigma_e^2 + \sigma_\omega^2 \tag{7}$$

## 2.2 Testing Site and Vehicle

The testing site is located at Texas A&M's Rellis campus in Bryan, Texas, a decommissioned Air Force base sprawling over 2000 acres. The campus has five major runways, most of which are used extensively by several different research bodies, not limited to transportation and autonomous vehicle studies. The campus is now home to several facilities dedicated to a wide variety of research areas, education, and technology development.

The vehicle used in this study is a 2014 Lincoln MKZ that was converted into an autonomous research platform using an aftermarket add-on kit. As a result, the MKZ's drive-by-wire (DBW) system is fully programmable and several built-in sensors provide information such as vehicle speed and yaw rate. Additional sensors include a pair of Velodyne Light Detection and Ranging (LIDAR) devices, a camera, a radar, a high-precision GPS, and an inertial measurement unit (IMU). A Linux-based rugged PC is used for interfacing with the vehicle. The primary means of message passing and communications is achieved through a robot operating system (ROS). In autonomous mode, the interfacing PC has full control over the vehicle; the user can exit this mode by taking control of the vehicle or pressing an emergency red switch located near the central console.

## 2.3 Experiment Design

To verify the effectiveness, robustness, sensitivity, and responsiveness of the DW technique in the context of AVs, a simple autonomous driving operation was carried out. The MKZ started at the origin and traveled along a circle at the specified speed (3 m/s) and yaw rate (0.5 rad/s), and the DW implemented in the controller was expected to detect the different attacks injected into the sensing system, specifically, into the yaw rate measurement.

### 2.3.1 Control Technique

To achieve the control goal of autonomous circling at the desired speed and yaw rate, proportional, integral, and derivative (PID) control, the most extensively used feedback control technique, was adopted. Speed error and yaw rate error (i.e., difference between desired and actual values) are the controller inputs, and throttle/brake pedal position and steering wheel angle are the controller outputs. The controller design consisted of tuning the PID gains and is omitted herein.

### 2.3.2 Attack Injection

To test the robustness of the DW technique, the following three sets of malicious attacks, each with three different parameters, were injected to tamper with the measurement of the yaw rate:

- Random noise with zero mean and variance ($rad^2/s^2$) of 2.5, 0.5, 0.1
- Constant bias and magnitude (rad/s) of 2.5, 0,5, 0.1
- Intermittent toggling (positive to negative) and duration (s) of 2.5, 0.5, 0.1

### 2.3.3 Watermark Sizes

For the exploration of the sensitivity and responsiveness of the DW technique (i.e., to answer the question of how small the watermark can be and still able to promptly detect attacks), three different sizes of watermarks were tested for each attack with each parameter. Superimposed on the steering wheel angle control, the watermark at each control step was a random number drawn from a Gaussian distribution with zero mean and a variance of 0.01, 0.001, and 0.0001 $rad^2/s^2$.

# 3. Calculations

## 3.1 System Modeling

Testing at low speeds, the widely used bicycle model can be employed to represent the dynamics of the vehicle. As shown in Fig. 1, the geometric relationship among the variables is clear. With counterclockwise being the positive direction, the yaw rate $\dot{\varphi}[t]$ is defined as

$$\dot{\phi}[t] = \frac{u[t]}{r[t]}, \tag{8}$$

where $u[t]$ and $r[t]$ are the vehicle's speed and turning radius, respectively.

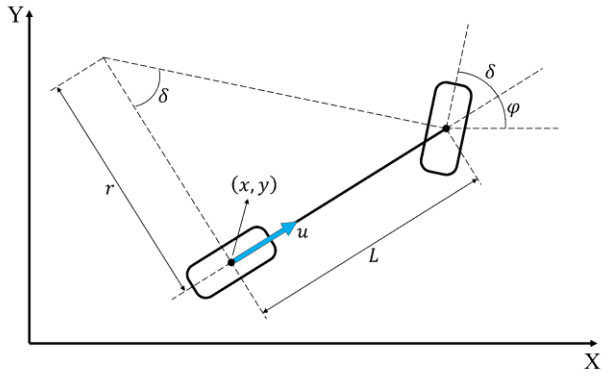


**Fig. 1 Bicycle model for the Lincoln MKZ**

With

$$\tan\delta[t] = \frac{L}{r[t]}, \tag{9}$$

the yaw rate can be expressed as

$$\dot{\phi}[t] = \frac{u[t]}{L}\tan\delta[t], \tag{10}$$

where $\delta[t]$ is the steering angle of the front wheel and $L$ is the wheelbase (i.e., distance between the front and rear axles). The linear relationship between $\delta[t]$ and the steering wheel angle $\theta[t]$ is

$$\delta[t] = a\theta[t], \tag{11}$$

where $a$ is the steering ratio ( $\frac{1}{14.8}$ for the MKZ). As discussed previously, $\theta[t]$ is calculated by the PID-based control function $f_c(\cdot)$, whose inputs are the measured yaw rate $\hat{\varphi}[t]$ and desired yaw rate $\varphi_d$. Thus, the actual steering angle applied is

$$\theta[t] = f_c\left(\phi_d, \hat{\phi}[t]\right) + n_{steer}[t], \tag{12}$$

where $n_{steer}[t]$ is the noise in the steering wheel control and $\hat{\varphi}[t]$ is the yaw rate measurement under different cyber attacks.

### 3.2 DW Tests Formulation

Given Eqs. 10–12, as well as the mechanism of the DW technique, the watermark $e_{DW}$ is superimposed on the nominal steering wheel angle control as follows.

Equation 13 can be reorganized into

$$\dot{\phi}[t] = \frac{u[t]}{L}\tan\left(a\left(f_c\left(\phi_d, \hat{\phi}[t]\right) + n_{steer}[t] + e_{DW}[t]\right)\right) \tag{13}$$

and

$$\arctan\left(\frac{L}{u[t]}\dot{\phi}[t]\right) = af_c\left(\phi_d, \hat{\phi}[t]\right) + an_{steer}[t] + ae_{DW}[t]. \tag{14}$$

Therefore, the two DW tests can be formulated as follows:

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\left(\arctan\left(\frac{L}{u[t]}\dot{\phi}[t]\right) - af_c\left(\phi_d, \hat{\phi}[t]\right)\right)^2 = a^2\left(\sigma_{steer}^2 + \sigma_{DW}^2\right) \tag{15}$$

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\left(\arctan\left(\frac{L}{u[t]}\dot{\phi}[t]\right) - af_c\left(\phi_d, \hat{\phi}[t]\right) - ae_{DW}[t]\right)^2 = a^2\sigma_{steer}^2, \tag{16}$$

where $\sigma_{steer}^2$ and $\sigma_{DW}^2$ are the variances of the steering actuation noise and dynamic watermark signal, respectively.

## 4. Results and Discussion

This section presents the experimental results, where the vehicle's yaw rate and the DW test results are plotted over time. For clarification, we offer the following:

- Test results are the absolute differences between the left sides and right sides of Eqs. 15 and 16, where closeness to 0 implies that the tests are passed and do not raise an alarm; whereas, if the absolute differences are significantly greater than 0 that implies that the attacks are detected.
- Each plot contains the results of the DW tests using three different watermark sizes for one single attack.
- In each plot, over the same time line with the DW test results, the associated yaw rate measurement with the largest watermark (i.e., 0.01 rad/s) is presented, from which one can tell that even the largest watermark did not

cause any noticeable fluctuation in the yaw rate, verifying that the DW hardly affected the nominal control performance.

- Data collected in the first second (i.e., 0 to 1 s) were cut off since the data logging may somehow be affected by the transitioning of the vehicle's state (from static to moving) as well as the initialization of the ROS programs.
- Because Tests 1 and 2 generated almost identical results in this study, only the Test 1 results are presented.
- All three types of cyber attacks (i.e., random noise, constant bias, and intermittent toggling) started at 10 s.

## 4.1 Random Noise Detection

First, the cyber attacks in the form of random noise with zero mean and different variances were injected. The yaw rate under attack and the DW test results are plotted in Figs. 2–4. It can be seen that when the variance is 2.5 or 0.5 $rad^2/s^2$, the test results increased immediately in response to the start of the attack at 10 s, from which point the yaw rate became rather noisy. The different watermark sizes (i.e., 0.01, 0.001, and 0.0001 $rad^2/s^2$) did not make much difference in the detection of the attacks. When the attack featured a small variance of 0.1 $rad^2/s^2$, resembling ambient noise, all three watermarks failed to detect the attacks as the test results barely grew.

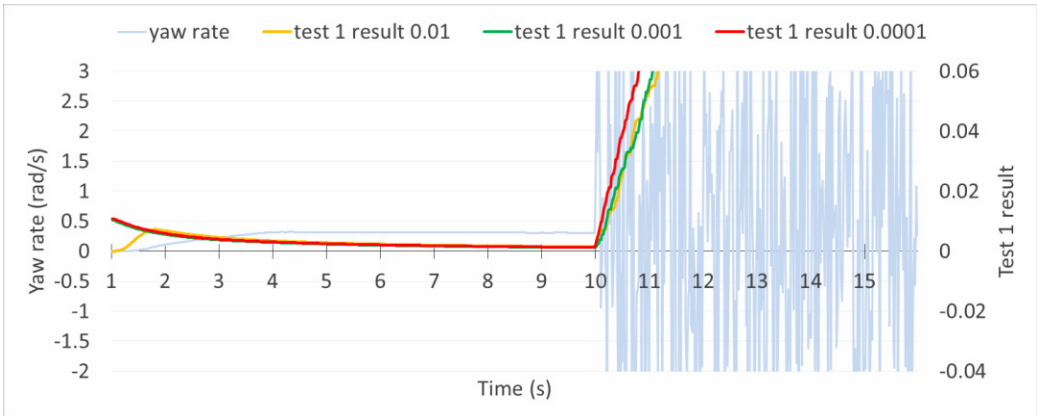


**Fig. 2 Detection of the random noise injection attack with a variance of 2.5 $rad^2/s^2$**

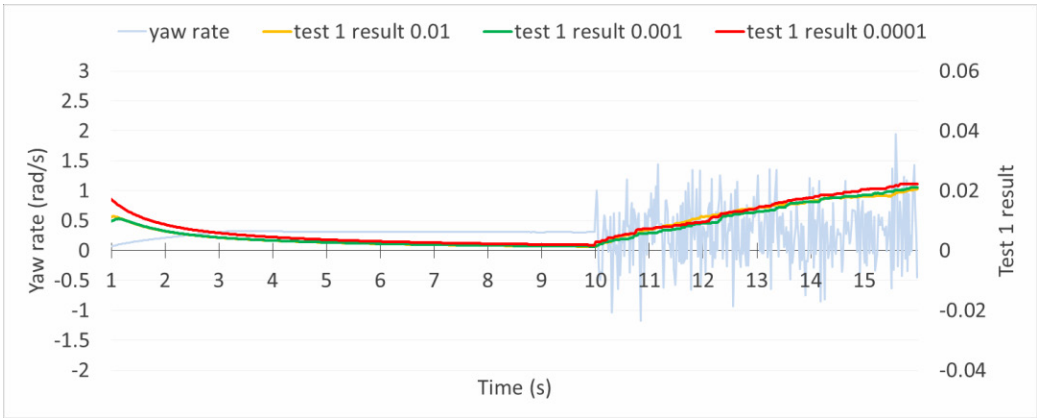


**Fig. 3 Detection of the random noise injection attack with a variance of 0.5 rad²/s²**

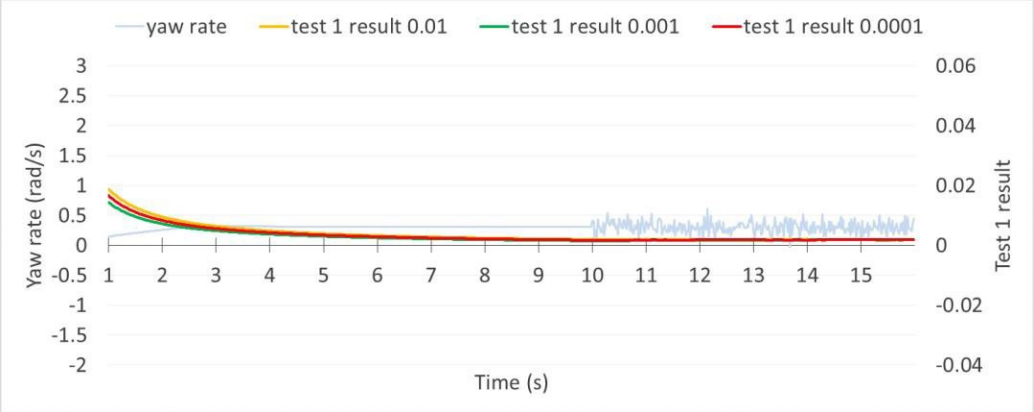


**Fig. 4 Detection of the random noise injection attack with a variance of 0.1 rad²/s²**

## 4.2 Constant Bias Detection

The second type of cyber attack studied was constant bias with different magnitudes. Shown in Figs. 5–7, when the attack started at 10 s, a constant bias with a magnitude of 2.5, 0.5, or 0.1 rad/s was added to the measurement of yaw rate. Similar to the previous case, the attacks with greater biases (2.5 and 0.5 rad/s) were easily caught by the DW technique with all three watermark sizes; whereas, the small constant bias (0.1 rad/s) escaped detection due to insignificant interference with the yaw rate measurement.

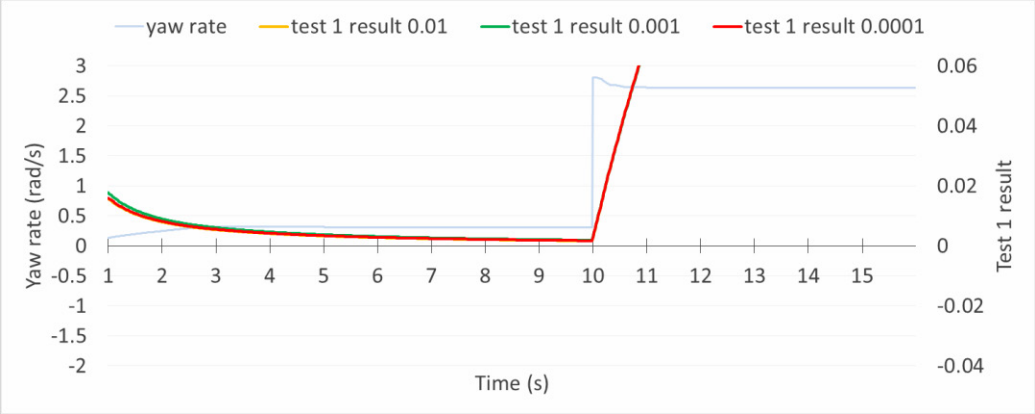


**Fig. 5 Detection of the constant bias attack with a magnitude of 2.5 rad/s**

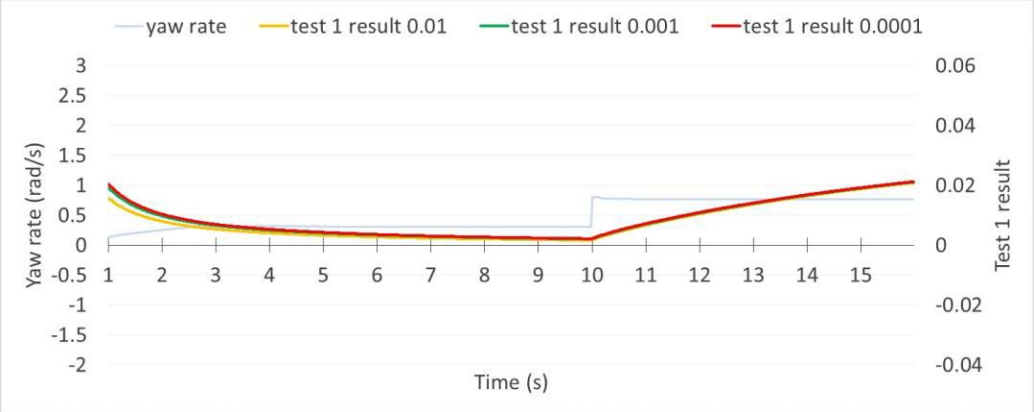


**Fig. 6 Detection of the constant bias attack with a magnitude of 0.5 rad/s**

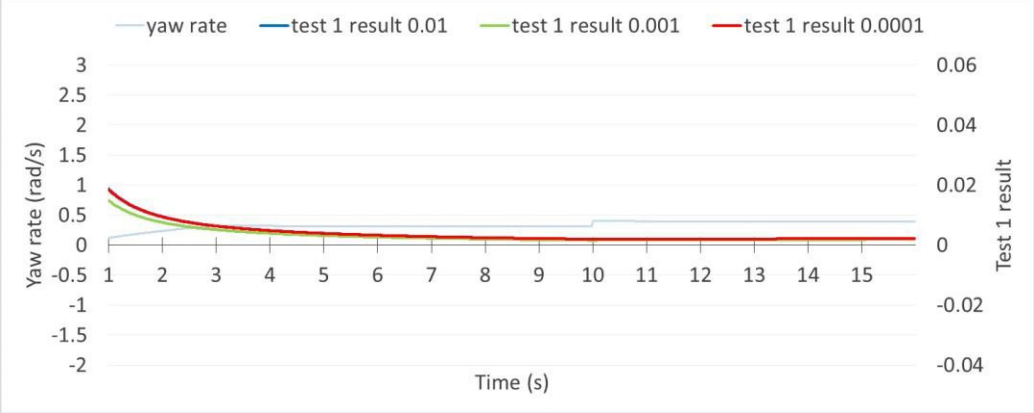


**Fig. 7 Detection of the constant bias attack with a magnitude of 0.1 rad/s**

## 4.3 Intermittent Toggling Detection

As the last set of attacks, the intermittent toggling was employed to tamper with the yaw rate measurement by "toggling" the measured value from positive to negative for 2.5, 0.5, and 0.1 s, as illustrated in Figs. 8–10, where the test signals increased rapidly during the toggling attacks. The longer the attack lasted, the greater the test signals increased, which implied the system was under attack. Nevertheless, when the duration of the toggling was short, such as 0.1 s in Fig. 10, the test signals might have not risen high enough to trigger failsafe actions.

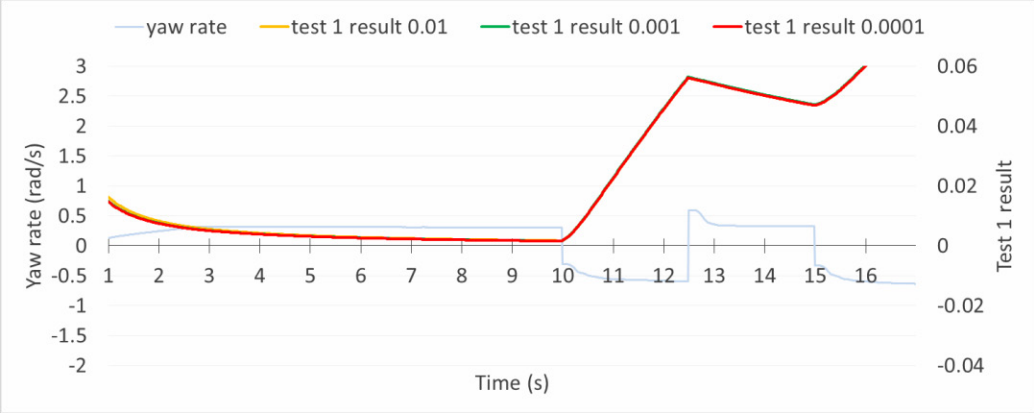


**Fig. 8 Detection of the intermittent toggling attack with a duration of 2.5 s**

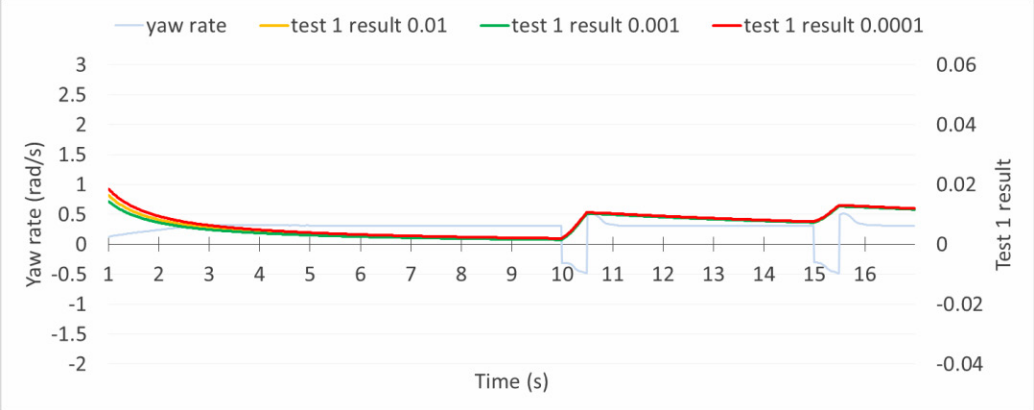


**Fig. 9 Detection of the intermittent toggling attack with a duration of 0.5**

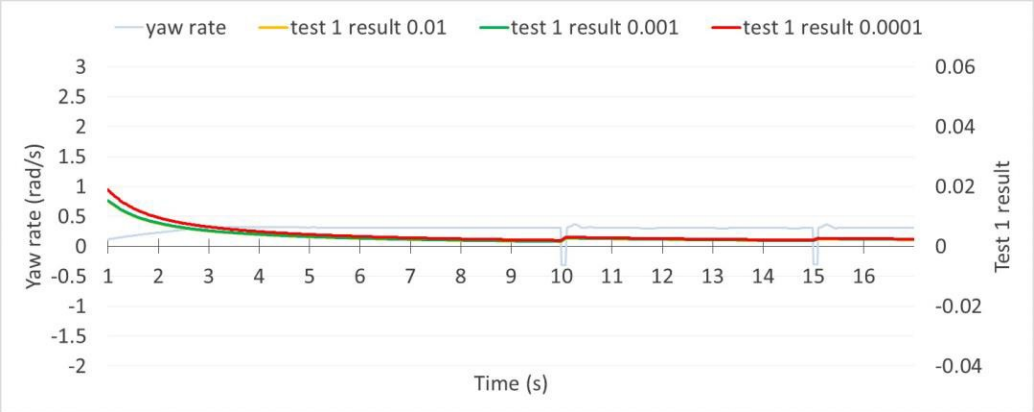


**Fig. 10 Detection of the intermittent toggling attack with a duration of 0.1 s**

## 4.4 Discussion

As demonstrated in the AV experiments, the DW technique was able to detect the three different types of cyber attacks (i.e., random noise injection attack, constant bias attack, and intermittent toggling attack) on the measurement of yaw rate in the control task of autonomous circling at a desired speed and yaw rate. Nevertheless, there are cases in which the DW technique failed to identify the attacks: when the impacts of the attacks on the system were too mild to be qualified as an "attack". Therefore, it is reasonable to declare the effectiveness and robustness of the DW technique.

In terms of sensitivity, the size of the watermark employed in this study was as low as 0.0001 rad/s, whose high performance was almost the same as those of 0.01 and 0.001 rad/s. The influence of the largest watermark (i.e., 0.01 rad/s) on the nominal yaw rate control was shown to be negligible in the last section; that is to say, extremely small watermarks with barely any impact on the nominal control (and no perceivable effect for the passengers) can be adopted and are still capable of detecting cyber attacks on the sensing system.

Another key feature of the DW technique is the fast responsiveness. As illustrated in the experimental results, for all types of attacks with all levels of severity, the DW test results increased immediately when the attack started, no matter how small the watermark was. That being said, setting the threshold for the test result is a tradeoff between false alarms and a time delay for triggering failsafe actions.

## 5. References


1. National Highway Traffic Safety Administration. Automated driving systems 2.0: a vision for safety. Washington (DC): Department of Transportation (US); 2017. Report No.: DOT HS 812 442.
2. Fagnant DJ, Kockelman K. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Trans Res Part A: Pol Prac. 2015;77:167–181.
3. Yăgdereli E, Gemci C, Aktaş AZ. A study on cyber-security of autonomous and unmanned vehicles. JDMS. 2015;12:369–381.
4. Axelrod CW. Cybersecurity in the age of autonomous vehicles, intelligent traffic controls and pervasive transportation networks. 2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT); 2017. IEEE. p. 1–6.
5. Onishi H, Wu K, Yoshida K, Kato T. Approaches for vehicle cybersecurity in the US. IJAE. 2017;8:1–6.
6. Parkinson S, Ward P, Wilson K, Miller J. Cyber threats facing autonomous and connected vehicles: future challenges. IEEE T ITS. 2017;18;2898–2915.
7. Satchidanandan B, Kumar PR. Dynamic watermarking: active defense of networked cyber–physical systems. Proc IEEE; 2016. Vol. 105. p. 219–240.
8. Huang T, Satchidanandan B, Kumar P, Xie L. An online detection framework for cyber attacks on automatic generation control. IEEE TPS. 2018;33:6816–6827.
9. Kim J, Ko WH, Kumar P. Cyber-security with dynamic watermarking for process control systems. AIChE Annual Meeting; 2019.
10. Ko WH, Satchidanandan B, Kumar P. Theory and implementation of dynamic watermarking for cybersecurity of advanced transportation systems. 2016 IEEE Conference on Communications and Network Security (CNS); 2016. p. 416–420.
11. Schenato L, Sinopoli B, Franceschetti M, Poolla K, Sastry SS. Foundations of control and estimation over lossy networks. Proc IEEE; 2007;95:163–187.
12. Sinopoli B, Schenato L, Franceschetti M, Poolla K, Sastry S. Optimal linear lqg control over lossy networks without packet acknowledgment. Asian J Control. 2008;10:3–13.

13. Gupta V, Hassibi B, Murray RM**.** Optimal LQG control across packetdropping links. Sys Con Lett. 2007;56:439–446.

14. Elia N, Mitter SK. Stabilization of linear systems with limited information. IEEETAC. 2001; 46:1384–1400.

15. Liu X, Goldsmith A. Kalman filtering with partial observation losses. 2004 43rd IEEE Conference on Decision and Control (CDC); 2004. p. 4180–4186. Vol. 4. IEEE Cat. No. 04CH37601.

16. Mo Y, Sinopoli B. Secure control against replay attacks. 2009 47th annual Allerton conference on communication, control, and computing (Allerton); 2009. IEEE. p. 911–918.

17. Mo Y, Chabukswar R, Sinopoli B. Detecting integrity attacks on SCADA systems. IEEE T CST. 2013;22:1396–1407.

18. Mo Y, Weerakkody S, Sinopoli B. Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs. IEEE CSM. 2015;35:93–109.

19. Weerakkody S, Mo Y, Sinopoli B. Detecting integrity attacks on control systems using robust physical watermarking. 53rd IEEE Conference on Decision and Control; 2014. p. 3757–3764.

20. Satchidanandan B, Kumar PR. On minimal tests of sensor veracity for dynamic watermarking-based defense of cyber-physical systems. 9th International Conference on Communication Systems and Networks (COMSNETS); 2017. IEEE. p. 23–30.

21. Satchidanandan B, Kumar PR. Secure control of networked cyberphysical systems. IEEE 55th Conference on Decision and Control (CDC); 2016. IEEE. p. 283–289.

22. B Satchidanandan B, Kumar PR. Defending cyber-physical systems from sensor attacks. International Conference on Communication Systems and Networks; 2017. Springer. p. 150–176.

23. Teixeira A, Shames I, Sandberg H, Johansson KH. Revealing stealthy attacks in control systems. 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton); 2012. IEEE. p. 1806–1813.

24. Gisdakis S, Giannetsos T, Papadimitratos P. Shield: A data verification framework for participatory sensing systems. Proc. 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks. 2015:16.

25. Cárdenas AA, Amin S, Sastry S. Research challenges for the security of control systems. HotSec; 2008.

26. Cui J, Sabaliauskaite G. On the alignment of safety and security for autonomous vehicles. IARIA Cyber; 2017; Barcelona, Spain.

27. Straub J, McMillan J, Yaniero B, Schumacher M, Almosalami A, Boatey K, Hartman J. Cybersecurity considerations for an interconnected self-driving car system of systems. 12th System of Systems Engineering Conference (SoSE); 2017. IEEE. p. 1–6.

## List of Symbols, Abbreviations, and Acronyms

| | |
|---|---|
| ARL | Army Research Laboratory |
| ARO | Army Research Office |
| AV | autonomous vehicle |
| CCDC | US Army Combat Capabilities Development Command |
| CPS | cyber-physical system |
| DBW | drive-by-wire |
| DoS | denial of service |
| DW | dynamic watermarking |
| FACT | failure, attack, and countermeasure |
| GNSS | Global Navigation Satellite System |
| GPS | global positioning system |
| i.i.d. | independent and identically distributed |
| IMU | inertial measurement unit |
| LIDAR | Light Detection and Ranging |
| PC | personal computer |
| PID | proportional, integral, and derivative |
| ROS | robot operating system |
| V2V | vehicle-to-vehicle |

| | |
|---|---|
| 1 (PDF) | DEFENSE TECHNICAL INFORMATION CTR DTIC OCA |
| 1 (PDF) | CCDC ARL FCDD RLD CL TECH LIB |
| 1 (PDF) | CCDC ARL FCDD RLW P S RAIO |